

METHODS AND SYSTEMS FOR OUTPUT OF SEARCH RESULTS

FIELD OF THE INVENTION

[0001] The invention generally relates to search engines. More particularly, the invention relates to methods and systems for output of search results.

BACKGROUND OF THE INVENTION

[0002] A search engine or search engine program is a widely used mechanism for allowing users to search vast numbers of documents for information. Automated general search engines locate documents, such as web pages, by matching terms from a user entered search query to an indexed corpus of web pages. A conventional network search engine, such as the Google™ search engine, returns a search result set in response to the search query submitted by the user. The search result set can comprise a ranked list of documents with a link to each document and a summary of the document can be returned to the user. The search engine can rank or sort the individual articles or documents in the result set based on a variety of measures, such as, the number of times the search terms appear in the document and the number of documents that contain a link to a document. For example, one known method, described in an article entitled “The Anatomy of a Large-Scale Hypertextual Search

Engine," by Sergey Brin and Lawrence Page, assigns a degree of importance to a document, such as a web page, based on the link structure of the web page.

[0003] Various documents, such as web pages, present items for sale. Some shopping documents allow users to purchase items, either directly, such as by clicking on a link, or indirectly, such as by providing a telephone number to complete a transaction. Users wishing to compare prices on an item from different vendors can enter a query for the item in a general search engine and obtain a list of relevant documents. Similarly, there may be different versions of the item and the user may desire to see which version each vendor carries. In order to compare prices or versions, the user must often visit several documents presenting the item for sale in conventional systems. Additionally, documents may be present in the search result set that are not shopping documents, but only discuss the item, such as reviews. Manually searching through a vast number of documents to extract attributes of the item can be extremely time consuming and impractical for a large number of documents.

SUMMARY

[0004] Embodiments of the present invention comprise methods and systems for output of search results. In one embodiment, a search engine implements a method comprising receiving a search query, identifying a plurality of item identifiers responsive to the search query, identifying a first group of item identifiers from the

plurality of item identifiers, wherein the first group of item identifiers was obtained by a first method, identifying a second group of items from the plurality of item identifiers, wherein the second group of item identifiers was obtained by a second method, and causing the output of all or a plurality of the item identifiers, comprising providing a cue to distinguish between the item identifiers from the first group and the item identifiers from the second group.

[0005] This exemplary embodiment is mentioned not to limit or define the invention, but to provide an example of an embodiment of the invention to aid understanding thereof. Exemplary embodiments are discussed in the Detailed Description, and further description of the invention is provided there. Advantages offered by the various embodiments of the present invention may be further understood by examining this specification.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] These and other features, aspects, and advantages of the present invention are better understood when the following Detailed Description is read with reference to the accompanying drawings, wherein:

[0007] Figure 1 is a diagram illustrating an exemplary environment in which one embodiment of the present invention may operate;

[0008] Figure 2 is a flow diagram illustrating a method of outputting search results to a user in one embodiment of the present invention; and

[0009] Figure 3 is a diagram of an exemplary search results page generated for display by a search engine, in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION OF SPECIFIC EMBODIMENTS

[0010] Embodiments of the present invention comprise methods and systems for output of search results. Referring now to the drawings in which like numerals indicate like elements throughout the several figures, Figure 1 is a block diagram illustrating an exemplary environment for implementation of one embodiment of the present invention. The system 100 shown in Figure 1 includes multiple client devices 102a-n in communication with a server device 104 over a network 106. The network 106 shown includes the Internet. In other embodiments, other networks, such as an intranet may be used. Moreover, methods according to the present invention may operate within a single computer.

[0011] The client devices 102a-n shown each includes a computer-readable medium, such as a random access memory (RAM) 108 coupled to a processor 110. The processor 110 executes computer-executable program instructions stored in memory 108. Such processors may include a microprocessor, an ASIC, and state machines. Such processors include, or may be in communication with, media, for example computer-readable media, which stores instructions that, when executed by the processor, cause the processor to perform the steps described herein.

Embodiments of computer-readable media include, but are not limited to, an electronic, optical, magnetic, or other storage or transmission device capable of providing a processor, such as the processor 110 of client 102a, with computer-readable instructions. Other examples of suitable media include, but are not limited to, a floppy disk, CD-ROM, DVD, magnetic disk, memory chip, ROM, RAM, an ASIC, a configured processor, all optical media, all magnetic tape or other magnetic media, or any other medium from which a computer processor can read instructions. Also, various other forms of computer-readable media may transmit or carry instructions to a computer, including a router, private or public network, or other transmission device or channel, both wired and wireless. The instructions may comprise code from any suitable computer-programming language, including, for example, C, C++, C#, Visual Basic, Java, Python, Perl, and JavaScript.

[0012] Client devices 102a-n may also include a number of external or internal devices such as a mouse, a CD-ROM, DVD, a keyboard, a display, or other input or output devices. Examples of client devices 102a-n are personal computers, digital assistants, personal digital assistants, cellular phones, mobile phones, smart phones, pagers, digital tablets, laptop computers, Internet appliances, and other processor-based devices. In general, a client device 102a may be any type of processor-based platform that is connected to a network 106 and that interacts with one or more application programs. Client devices 102a-n may operate on any suitable operating system capable of supporting a browser or browser-enabled application, such as

Microsoft® Windows® or Linux. The client devices 102a-n shown include, for example, personal computers executing a browser application program such as Microsoft Corporation's Internet Explorer™, Netscape Communication Corporation's Netscape Navigator™, and Apple Computer, Inc.'s Safari™.

[0013] Using the client devices 102a-n, users 112a-n can communicate over the network 106 with each other and with other systems and devices coupled to the network 106. As shown in Figure 1, a server device 104 is also coupled to the network 106. In the embodiment shown, a user 112a-n generates a search query 114 at a client device 102a. The client device 102a transmits the query 114 to the server device 104 via the network 106. For example, a user 112a types a textual search query regarding an item into a query field of a web page of a shopping search engine interface or other client-side software displayed on the client device 102a, which is then transmitted via the network 106 to the server device 104. In the embodiment shown, a user 112a inputs a search query 114 at a client device 102a, which transmits an associated search query signal 122 reflecting the search query 114 to the server device 104. The search query 114 may be transmitted directly to the server device 104 as shown. In another embodiment, the query signal 122 may instead be sent to a proxy server (not shown), which then transmits the query signal 122 to server device 104. Other configurations are possible.

[0014] The server device 104 shown includes a server executing a shopping search engine application program, such as the Froogle™ search engine. Similar to

the client devices 102a-n, the server device 104 shown includes a processor 116 coupled to a computer-readable memory 118. Server device 104, depicted as a single computer system, may be implemented as a network of computer processors. Examples of a server device 104 are servers, mainframe computers, networked computers, a processor-based device, and similar types of systems and devices. Client processor 110 and the server processor 116 can be any of a number of computer processors, such as processors from Intel Corporation of Santa Clara, California and Motorola Corporation of Schaumburg, Illinois.

[0015] Memory 118 contains the shopping search engine application program, also known as a search engine 120. The search engine 120 locates relevant items by identifying item identifiers in response to a search query 114 from a user 112a-n. Items can include, for example, products, goods, services, and any other thing offered for sale. The search engine 120 then provides the result set 124 to the client 102a via the network 106. The result set 124 can comprise a ranked list of item identifiers and attributes for each item, such as a price, an image, an item description and an item version. The item identifier can be an item name, an item number, or link to an article offering the item for sale or any combination of item name, item number and link. The link can be a uniform resource locator (URL), a file name, an icon or anything else that identifies an article offering the item for sale. Information regarding items, comprising item identifiers and attributes associated with items, can be obtained by the search engine 120 by a variety of methods. Information regarding items can come

directly from vendors offering the items for sale in the form of a vendor feed prior to receiving a search query and can be stored in memory 118 or in a database 140 or other storage device accessible by the server device 104. Information regarding items can also be automatically extracted from articles offering the items for sale prior to receiving a search query and stored in the database 140. Information regarding items can also be automatically extracted from articles offering the items for sale after receipt of the search query.

[0016] Information on items can be extracted from articles prior to the search query by various systems and methods, including, for example, the systems and methods using a template-based information extraction approach described in U.S. Patent Application Serial No. 10/675,756 filed September 30, 2003, which is hereby incorporated herein by this reference. Additionally, information regarding items relevant to the search query 114 can be extracted from articles by using the search query, such as, for example, described in U.S. Patent Application Serial No. 10/731,916 [Attorney Docket No. GP-078-04-US, 53051/293400] filed December 10, 2003, which is hereby incorporated herein by this reference.

[0017] The search engine 120 can cause the output, such as the display, of the search result set 124 on the client device 102a. The item identifiers can be output with a cue that distinguishes between the item identifiers based on the method by which item information was obtained or a degree of certainty of the accuracy of the item information. For example, the items information received directly from a vendor

can be grouped together and displayed in a manner visually distinguishing these item identifiers from other item identifiers in the result set. Item information automatically extracted from articles can also be grouped together and displayed in a manner visually distinguishing these item identifiers from other item identifiers in the result set. In such an embodiment, the item identifiers from each group can be presented in the result set 124 to the user as visually distinct groups or the item identifiers from one group or each group can be displayed with a similar visual indication, such as an icon. Other ways of displaying the item identifiers based on the methods by which the item identifiers were initially obtained or the degree of certainty of the accuracy of the information provided with each item identifiers can be used. Systems and methods of displaying search results to a user are described in U.S. Patent Application Serial No. 10/673,957 filed September 29, 2003, which is hereby incorporated by this reference.

[0018] The search engine 120 can also comprise a document locator 130 and an attribute processor 132. In the embodiment shown, each comprises computer code residing in memory 118. In the embodiment shown, the server device 104, or related device, performs a crawl of the network 106 to locate articles, such as web pages, stored at other devices or systems connected to the network 106, and indexes the articles in memory 118 or on another data storage device prior to receiving and processing a search query 114. Articles include, for example, web pages of various formats, such as HTML, XML, XHTML, Portable Document Format (PDF) files, and

word processor, database, and application program document files, audio, video, or any other documents or information of any type whatsoever made available on a network (such as the Internet), a personal computer, or other computing or storage means. The embodiments described herein are described generally in relation to HTML files or documents, but embodiments may operate on any type of article, including any type of image. During or after the crawl of the network the server device 104 or other device can determine if any of the documents crawled are shopping documents. A shopping document is a document that presents items for sale and in which the items can be purchased through interaction with the document or related documents.

[0019] In response to a search query signal 122, the document locator 130 can identify items by item identifiers and/or item attributes previously stored in the database 140 and shopping documents that contain items relevant to the search query signal 122. A shopping document can contain a single item or multiple items relevant to the search query. The attribute processor 132 can identify and extract certain attributes associated with the item or items relevant to the search query from shopping documents. In one embodiment, the attribute processor 132 identifies and extracts the price of the item and an image associated with the item from the documents located by the document locator 130 that are relevant to the search query.

[0020] It should be noted that the present invention may comprise systems having a different architecture than that which is shown in Figure 1. For example, in

some systems according to the present invention, the attribute processor 132 may be located external to the search engine 120 and may perform its functions offline before a search query from a user. The system 100 shown in Figure 1 is merely exemplary, and is used to explain the exemplary method shown in Figure 2.

[0021] Various methods in accordance with the present invention may be carried out. For example, in one embodiment, a search query is received, a plurality of item identifiers responsive to the search query are identified, a first group of item identifiers, obtained by a first method, from the plurality of item identifiers is identified, a second group of item identifiers, obtained by a second method, from the plurality of item identifiers is identified, and the item identifiers are output such that a cue is provided to distinguish between the item identifiers from the first group and the item identifiers from the second group. In another embodiment, a third group of item identifiers, obtained by a third method, from the plurality of item identifiers is identified, and the item identifiers are output such that a cue is provided to distinguish between the third group of item identifiers and the item identifiers from the first and second groups.

[0022] In one embodiment, the first method can be vendor feeds or automatic extraction of item identifiers and associated attributes from shopping documents. The extraction can be performed at least in part by a template-based extraction method. The extraction can also be based at least in part on the search query. In one embodiment, the first method can be vendor feeds, the second method can be

template-based extraction of item identifiers and associated attributes from shopping documents, and the third method can be extraction of item information from shopping document using the search query.

[0023] In one embodiment, each item identifier in the first group is output with a visual indication, such as an icon, making it visually distinct from the other item identifiers when displayed. The output of the first group of item identifiers can be a first list and the output of the second group of item identifiers can be a second list, wherein the first list can be visually separated from the second list when displayed. Alternatively, the output of the first group of item identifiers can be a first grid and the output of the second group of item identifiers can be a second grid, wherein the first grid can be visually separated from the second grid when displayed.

[0024] In another embodiment, a search query for an item identifier received, a plurality of item identifiers responsive to the search query are identified, a first group of item identifiers, having a first degree of certainty, from the plurality of item identifiers are identified, a second group of item identifiers, having a second degree of certainty, from the plurality of item identifiers are identified, and the item identifiers are output such that a cue is provided to distinguish between the item identifiers from the first group and the item identifiers of the second group. The degree of certainty can be related to the method of obtaining item information or can relate to another factor that can influence the accuracy of the information provided for each item identifier.

[0025] Figure 2 illustrates an exemplary method 200 that provides a method for the output of search results. This exemplary method is provided by way of example, as there are a variety of ways to carry out methods according to the present invention. The method 200 shown in Figure 2 can be executed or otherwise performed by any of various systems. The method 200 is described below as carried out by the system 100 shown in Figure 1 by way of example, and various elements of the system 100 are referenced in explaining the example method of Figure 2.

[0026] In 202, the search engine 120 receives a search query signal 122 from client device 102a via the network 106. A user 112a can input text representing an item the user is interested in into a client device 102a. The client device 102a can create an associated search query signal 122 representative of the search query 114 input by the user and transmit the search query signal 122 to the server device 104 via the network 106. For example, if the user 112a is interested in purchasing a camera, the user can input "camera" in the client device 102a and the client device can send an associated search query signal 122 to the server device 104.

[0027] In 204, the document locator 130 identifies item identifiers identifying items relevant to the item requested in the search query. For example, if the search query is "camera", the document locator 130 identifies cameras as items relevant to the query. Item information, comprising item identifiers and item attributes, can be received by the search engine 120 and indexed and stored in the database 140 prior to receiving the search query. This item information can be obtained directly from the

vendor of the items in the form of vendor feeds. For example, a vendor feed can be a data stream from the vendor and can include the item identifiers that the vendor is offering for sale and attributes associated with each item, such as a price of the item, and an image of the item. Item information can also be received prior to receiving the search query by the search engine 120 by extracting information from shopping documents using a template-based extraction method or other method. Item information can also be extracted from shopping documents by the search engine 120 by using the terms of the search query.

[0028] In 206, the search engine 120 can rank and sort the item identifiers relevant to the search query to form a result set. The search engine can rank the item identifiers by any known technique, such as a hyperlink structure-based ranking method. The search engine 120 can also determine the method used to obtain each item identifier and a degree of certainty of accuracy of the attributes associated with each item identifier. For example, the item information that was received through vendor feeds can have a high degree of certainty and can be grouped together and the item information received through extraction methods can have a lower degree of certainty and can be grouped together. In another embodiment, item information that is received through extraction methods can be further grouped by the particular extraction method used.

[0029] In 208, the search engine 120 causes the output, such as a display, of the search result set 124. The item identifiers can be output so that they are displayed

on a single document, such as a web page, or on multiple documents. In one embodiment, the search engine 120 can cause the display or other output of the item identifiers such as to distinguish between the item identifiers obtained by various methods through the use of a cue. A cue can comprise using different parts of a page, using different formatting, color, or font, using a header, or using different icons. For example, item identifiers from a first group obtained by vendor feeds can be displayed in list form separate from item identifiers from a second group obtained through extraction methods. Alternatively, the item identifiers from the first group or the second group or both can be displayed with a visual indication, such as an icon, indicating the method by which they were obtained. In one embodiment, the item identifiers obtained using a particular method may be formatted for display or other output differently—and thus themselves be visually distinct—relative to item identifiers obtained using another method (e.g., a different font, size, color or other attribute of the items may be used to help distinguish between items obtained using one method versus another). In another embodiment, the search engine 120 can cause the output of the item identifiers so as to provide a visual indication of the degree of certainty of the accuracy of the information associated with each item identifier. For example, item identifiers with a high degree of certainty can be displayed separately from item identifiers with a lower degree of certainty, such as two separate lists. Alternatively, item identifiers with a high degree of certainty or a lower degree of

certainty or both can be displayed each with a visual indication, such as an icon, indicating the associated degree of certainty.

[0030] Figure 3 illustrates an exemplary search results page 300 generated for display by the search engine 120. The search results page 300 displays a partial list view of search results 302 A-C and 304 A-B generated in response to the search query “Canon S400”. Each result 302 A-C and 304 A-B shown on the page 300 has an item identifier 310, 320, 322, 324, and 326 corresponding to an item determined by the search engine 120 to be relevant to the search query and a price and an image of the item. For example, result 302 comprises item identifier 310, a price 312, and an image 314. In the embodiment shown in Figure 3, the item identifier 310 comprises the item name and acts as a link to the shopping document offering the item for sale. The item identifier 310, in the embodiment shown in Figure 3, also comprises additional information regarding the purchase of the item – “w/ Free Tripod and Bag”.

[0031] The item identifiers 310, 320, 322, 324, and 326 are grouped by the method that the item information associated with each item identifier 310, 320, 322, 324, and 326 was initially obtained. In the embodiment shown, the item identifiers 310, 320, and 322 were obtained from vendor feeds. For example, the vendors of item identifiers item identifiers 310, 320, and 322 sent information about the items directly to the search engine 120 or related device and the search engine 120 or related device indexed and stored this information in the database 140 before the

search query was received. In processing the search query, the document locator 130 matched these results 302 A-C with the search query and returned them as relevant to the search query.

[0032] In the embodiment shown, the item information associated with item identifiers 324 and 326 was obtained by extracting the item identifiers and associated attributes from shopping documents. For example, the information contained in result 304 A could have been extracted from a shopping document by using a template-based extraction method and stored in the database 140 by the search engine 120 or related device 120 before the search query was received. The information contained in result 304 B could have been located and extracted by the attribute processor 132 by using the search query.

[0033] Text 306 separates the two groups of search results. The results can be grouped as shown in Figure 3 because the information received directly from the vendor generally has a higher degree of certainty than the information extracted from shopping documents and it is thus desirable to display more reliable information first to a user.

[0034] In the embodiment shown in Figure 3, the first group of item identifiers 310, 320, and 322 and the second group of item identifiers 324 and 326 are displayed in list form and the two lists are separated by the text 306. Alternatively, the two groups of item identifiers 310, 320, 322, 324, and 326 can both be displayed in grid

form and the two grids can be separated visually such as by a line or some other means. In another embodiment, each item identifier 310, 320, 322, 324, and 326, when displayed, can indicate the respective group that the item identifiers 310, 320, 322, 324, and 326 belongs to, such as, for example, by some visual indication displayed with each item.

[0035] In one embodiment, the second group of item identifiers 324 and 326 can be separated into two groups based on the method of extraction used to obtain the item. The display of these two groups can be separated visually.

[0036] It should be appreciated from the foregoing description of exemplary embodiments of the invention that numerous modifications are possible in other embodiments. For example, the invention could be used with various types of search mechanisms, databases and item identifiers (not just vendor-related), such as news, people/social networks, classifieds, etc. Furthermore, while exemplary methods for distinguishing between item identifiers obtained using various methods have been described, those skilled in the art will recognize from this description that various other methods may be employed without departing from the spirit and scope of the invention.

[0037] While the above description contains many specifics, these specifics should not be construed as limitations on the scope of the invention, but merely as exemplifications of the disclosed embodiments. Those skilled in the art will envision many other possible variations that are within the scope of the invention. The terms

first and second are used herein merely to differentiate one item identifier, group, method, article, or attribute from another item identifier, group, method, article, or attribute. The terms first and second are not used to indicate first or second in time, or first or second in a list, or other order, unless explicitly noted. For example, the "second" may come in time or in a list before the "first," unless it is otherwise explicitly indicated.